



SNS COLLEGE OF TECHNOLOGY

Coimbatore - 35



23BAE734 – Visual Analytic and Storytelling

Unit II – Tools and Techniques



Presented by

Ms.S.D.Shamini
Design Thinker

Redesigning Common Mind & Business Towards Excellence

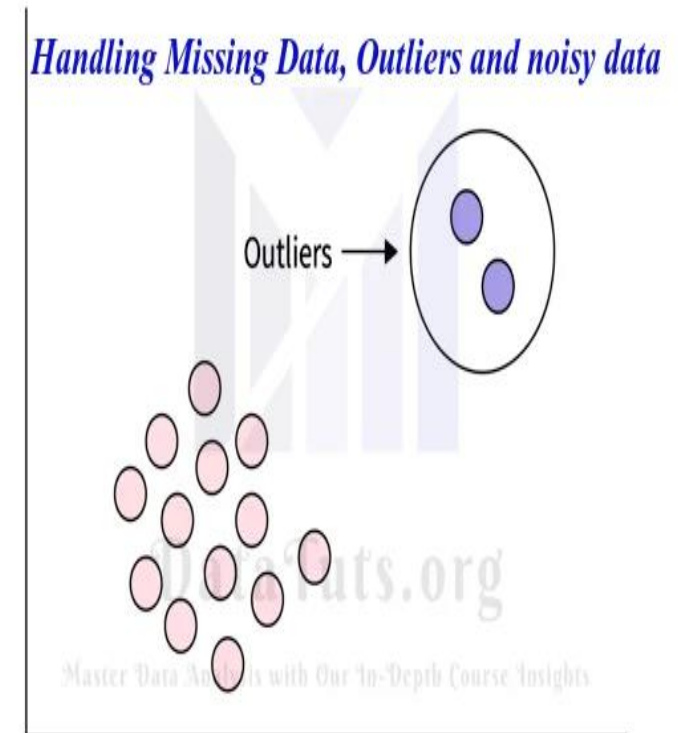


Build an Entrepreneurial Mindset Through Our Design Thinking FrameWork



Guess the Topic!!!

Handling Missing Data and Outliers





Recap



Missing Data

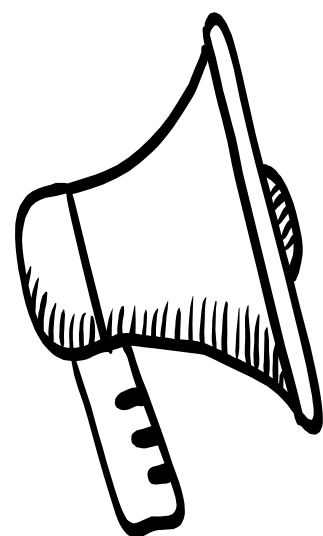
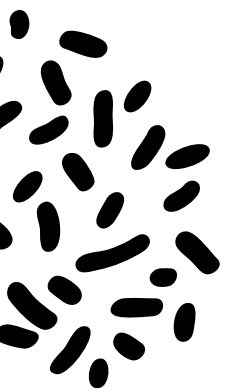
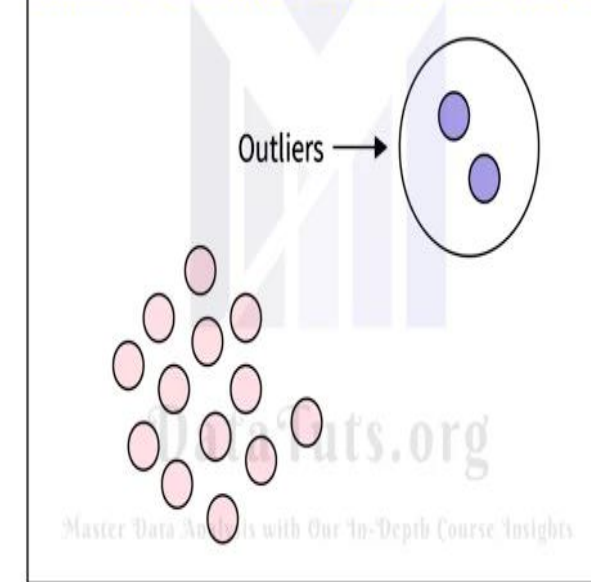
Types of Missing Data

Strategies for Handling Missing Data

Outliers

Handling outliers

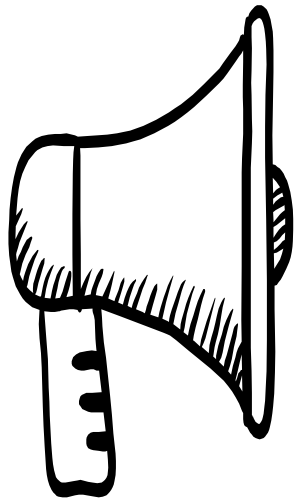
Handling Missing Data, Outliers and noisy data





Discussion about....

- ❑ Understanding Missing Data Mechanism
- ❑ Identifying and Visualizing Missing Data Patterns
- ❑ Imputation Techniques for Missing Data in Power BI
- ❑ Detecting Outliers Using Power BI Features
- ❑ Handling Outliers: Removal vs. Transformation
- ❑ Visualizing Data After Handling Missing Values and Outliers





1.133	53.66	51.09	
7222	52.75	7225	57555
8150	89.27	75.05	22455
42.95	17.88	89.09	25445
81.85	02.05	03.95	35399

Understanding Missing Data Mechanisms

1 Missing Completely at Random (MCAR)

The missing data is completely independent of any other data in the dataset.

2 Missing at Random (MAR)

The missing data is related to other observed variables in the dataset.

3 Missing Not at Random (MNAR)

The missing data is related to the missing value itself.



Identifying and Visualizing Missing Data Patterns



Technique

Description

Missing Value Analysis

Identifies the percentage of missing values in each column.

Heatmap

Visually represents missing values across the dataset.



Imputation Techniques for Missing Data in Power BI

Mean Imputation

Replaces missing values with the average of existing values.

Median Imputation

Replaces missing values with the median of existing values.

Mode Imputation

Replaces missing values with the most frequent value.



Detecting Outliers Using Power BI Features

Visual Inspection

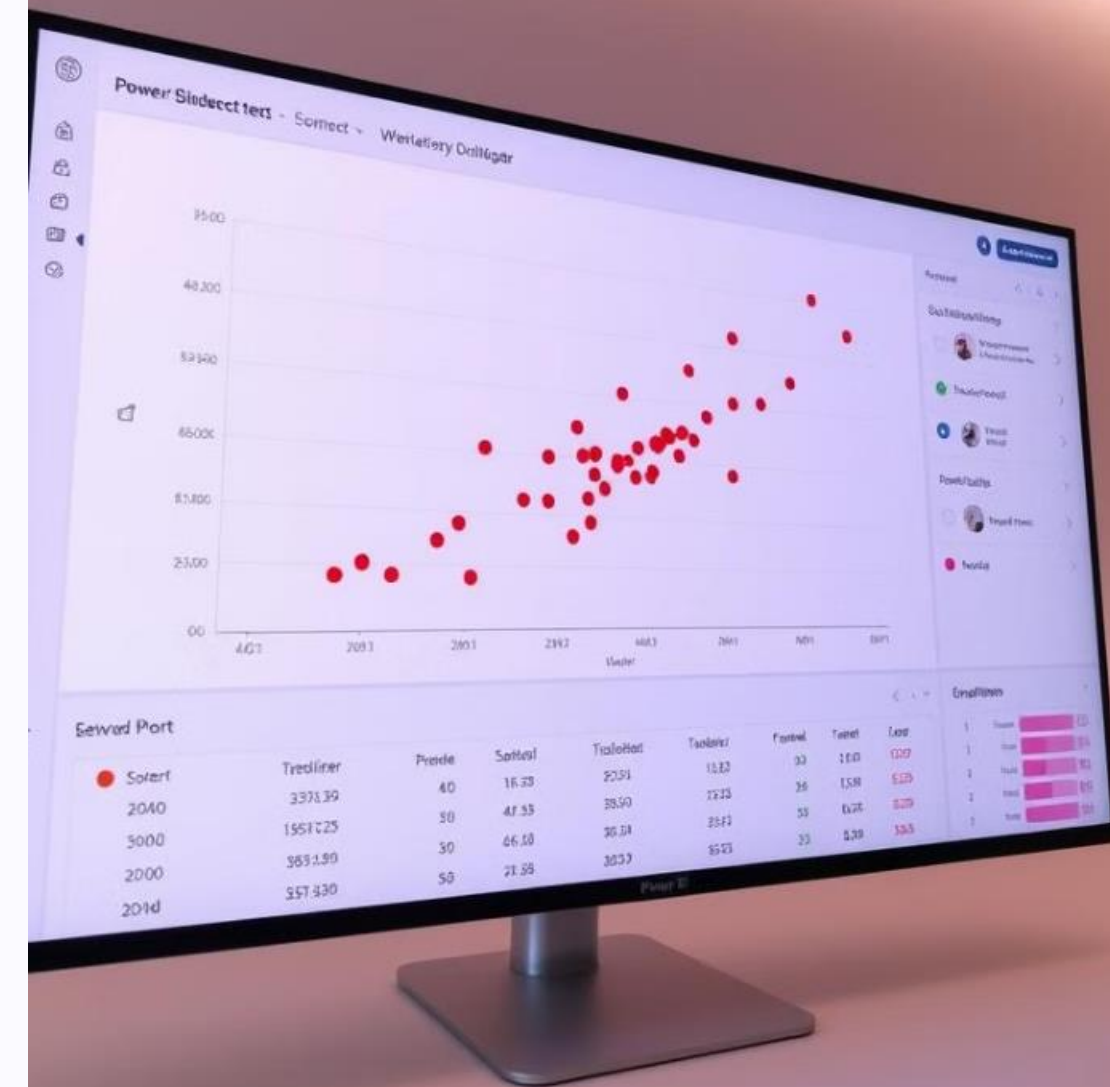
Use charts and graphs to identify data points that deviate significantly from the general trend.

Statistical Measures

Calculate statistical measures like standard deviation or interquartile range to identify outliers.

Power BI Features

Use Power BI features like "Anomaly Detection" to automatically detect outliers.





Handling Outliers: Removal vs. Transformation

1

Removal

Completely remove outliers from the dataset.

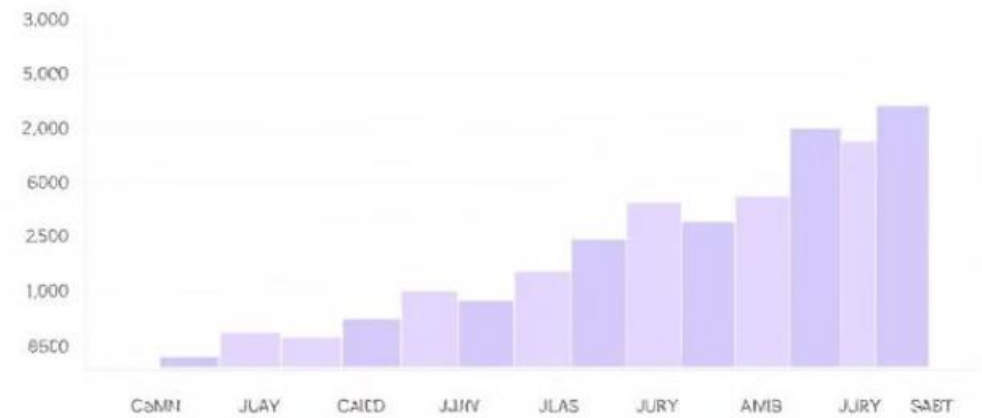
2

Transformation

Transform outliers to bring them closer to the rest of the data.



Visualizing Data After Handling Missing Values and Outliers



Bar Charts

Illustrate the distribution of data categories.



Scatter Plots

Show the relationship between two variables.

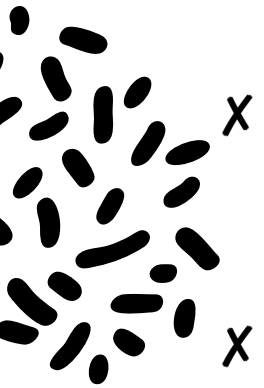


Histograms

Represent the frequency distribution of data.



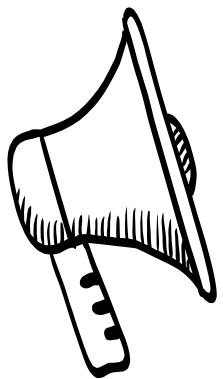
Knowledge Check



Which of the following techniques is commonly used to handle missing data in a dataset before visualization?

- A. Remove all missing data
- B. Replace missing values with the mean, median, or mode
- C. Ignore missing data and proceed
- D. Delete the entire dataset

x B) Replace missing values with the mean, median or mode





Summary



- ❑ Understanding Missing Data Mechanism
- ❑ Identifying and Visualizing Missing Data Patterns
- ❑ Imputation Techniques for Missing Data in Power BI
- ❑ Detecting Outliers Using Power BI Features
- ❑ Handling Outliers: Removal vs. Transformation
- ❑ Visualizing Data After Handling Missing Values and Outliers

I just need
the main ideas





References

- <https://www.linkedin.com/advice/3/how-can-you-identify-outliers-missing-values-your-dj0nc>



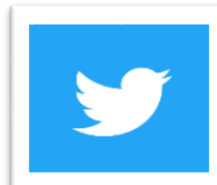
REACH US



snsinstitutions



snsinstitutions



snsinstitutions



snsinstitutions



snsinstitutions

Thanks!

